

ABSTRACT OF THE DISCLOSURE

This disclosure teaches a fast-computerized method for finding new repeating sequences and fragments via Region Definition and Transition Identification Procedure. New Repeating Sequences can be recognized when an unknown Query Sequence is compared and aligned with a plurality of previously stored sequence fragments. Using a Region Definition Procedure, each of the aligned sequences has a beginning and an end point that defines a region that is compared directly with the Query Sequence during the alignment process. A Transition Identification algorithm then recognizes different patterns of hits in the region transitions and detects new repeating sequences. Newly recognized repeating sequences are stored in a REP FILE for future use in identifying and masking repeat sequences found in new Query Sequences.